

WORKING P A P E R

An Educator Looks at Quality Improvement in Health Care

BRIAN STECHER

WR-241-EDU

February, 2005

This product is part of the RAND Education working paper series. RAND working papers are intended to share researchers' latest findings and to solicit informal peer review. They have been approved for circulation by RAND Education but have not been formally edited or peer reviewed. Unless otherwise indicated, working papers can be quoted and cited without permission of the author, provided the source is clearly referred to as a working paper. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors. RAND® is a registered trademark.

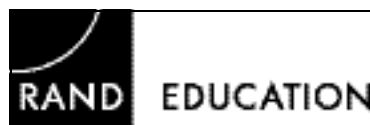


Table of Contents

1. Introduction.....	1
2. Accountability in Education.....	3
3. Quality Improvement in Health Care.....	5
Measuring Quality Based on Structure	5
Measuring Quality Based on Outcomes.	6
Risk Adjustment Methods.....	6
Measuring Quality Based on Process	8
Clinical Practice Guidelines.....	8
Quality Indicators.....	10
4. Process Measurement Applied to Education	15
Caveats.....	16
References	19

Figures and Tables

Figure 3.1. Prevention and Management of Diabetes Complications	9
Figure 3.2. Recommended Quality Indicators for Diabetes Mellitus	12
Table 4.1. Relevant Differences between Health Care and Education.....	17

Preface

This article is derived in whole or part from work sponsored by the U.S. government. Any opinions, findings, conclusions, or recommendations expressed in this article are those of the author(s) and do not necessarily reflect the views of the U.S. government or the National Academy of Sciences.

1. Introduction

It is not obvious that a paper on quality improvement in health care is relevant to a workshop on incentives and test-based accountability in education. Yet, I hope to convince you that it is wise to broaden the perspective of this workshop to encompass quality improvement efforts, not just accountability, and to broaden the scope to include other sectors, not just education. Health care offers an example of a system for judging the quality of practice directly. This paper argues that educators and educational policymakers should be trying to develop similar systems for education.

The following story encapsulates the issues that are elaborated in the paper. The encounter is fictitious, but all the important elements are real.

Some months ago, I met with Sarah, a newly hired researcher, who was visiting the RAND office in Santa Monica prior to moving here from Michigan. She had come to Santa Monica to fill out official papers, meet other staff, and learn about projects. After we finished talking about work, she asked for some advice about other relocation issues. She had to choose a health plan and wanted to know which would be the best choice, given the fact that her husband had diabetes and needed regular care to manage his disease. Although this is not my area of expertise, I knew enough to go to the Web and find the CAHPS (Consumer Assessment of Health Plans) survey results or the HEDIS (Health Plan Employer Data and Information Set) ratings of the available health plans based on general-purpose criteria. Not surprisingly, the three that RAND makes available were similarly rated in terms of overall quality. So, I looked further and found the results of a recent Quality Index study, which had specific information about treatment of patients with Diabetes Mellitus. The study showed that the typical diabetic patient in Southern California received only 45 percent of the care that was recommended by established clinical practice guideline. Fortunately, the three RAND providers exceeded the average, and one was notably higher, providing about 80 percent of the recommended care. Sarah was very grateful for the information, and said she would select that health plan.

Two or three months later, I saw Sarah in the cafeteria. "Thank you for the information," she said. "My husband is pleased with his new doctor." Her husband had recently been to the clinic for an initial visit. They did a thorough workup, and they seemed to be very competent. "They recommended that he change his diet," she explained, "and they want him to try some new medication I never heard of." I suggested that we look up the treatment recommendations. So, we looked on the Web for the National Guideline Clearinghouse and downloaded the latest clinical practice guidelines for diabetes. Sure enough, she recognized the name of the medication among those recommended for diabetics with particular co-existing conditions. She was pleased, and I was satisfied.

After she left, I pondered this whole exchange. If a patient can find out how well a clinic or hospital meets the standards for care for a given condition, why can't a parent find out how well a school is delivering instruction in reading? If a patient can obtain detailed

guidelines for optimum medical care for a given condition, why can't a parent find out what the recommended instructional activities are for a given subject and grade level? Why is health care addressing quality much more directly than education?

This paper is motivated by the questions raised in the vignette. I will argue that educational quality is not likely to improve if we just find the right outcome measure to indicate the performance of students at a particular point in time. Those outcomes are, at best, an indirect indicator of the quality of the curriculum and the teacher-student interaction. As many have observed, league tables of test results tell you more about the home environment of the students than the quality of instruction. Nor is quality likely to be enhanced by finding the right incentive to motivate teacher and administrator behavior. For the most part, teachers and administrators are working hard to produce better outcomes, but they don't know what behaviors they should be striving to enact. One third-grade teacher assigns a lot of homework, another does not. One elementary school groups student by ability, another groups students heterogeneously. Which is better? This paper argues that educators should be figuring out how to measure the core processes for which they are responsible and how to use these to drive improvement efforts.

2. Accountability in Education

In 2001, Congress passed the No Child Left Behind (NCLB), with the laudable goal that all students would be proficient in reading and mathematics by 2012. NCLB mandates a system of standards-based accountability, which is supposed to insure success at meeting this goal.

Specifically, states are required to implement the following components:

- Explicit content standards in reading and mathematics for students in grades 3-8 (and by 2005-06 in science for three grade ranges)
- Aligned annual student assessments that provide individual scores
- Performance standards designating basic, proficient and advanced levels
- Increasing targets for adequate yearly progress (AYP) for districts and schools
- Rewards and sanctions for districts and schools based on yearly performance
- Parental choice to select a different school or to request supplemental educational services from outside providers if the original school does not meet its target repeatedly.

This approach to educational improvement has been called “steering by results” (California Department of Education, 1998), a phrase that captures the idea that outcomes should be the focal point for educational planning. By sending clear signals about student performance and attaching real consequence to results, NCLB is supposed to succeed where previous reform efforts have failed.

Yet, many have pointed out the shortcomings of this approach. NCLB seems to make the implicit assumption that poor student performance is due to lack of information on student achievement or to lack of motivation on the part of teachers and administrators. For the NCLB remedies to be successful either:

- (a) Educators do not know enough about student performance, and standards-based annual testing results will provide the information they need to improve practice; or
- (b) Educators do not work hard enough at improvement, and the new sanctions and market mechanisms will motivate them to better their practice.

I am not convinced that either of these assumptions is valid. While I cannot prove this point empirically, it is certainly true that educators have had information about low student achievement for decades through state tests, national metrics and international comparisons. It is also true that educators who continue to work in the schools despite the challenges they face are highly motivated individuals. It seems far more likely to me that educators do not know how to change their practice to raise student performance. (It may be the case that other pressures constrain their effectiveness or that other conditions stymie their efforts; for now I include these under the general problem of lack of know-how.)

The underlying problem is neither information nor motivation, but lack of capacity to improve student learning (or address countervailing pressures or conditions). Teachers and administrators have exhausted their existing repertoire of curriculum, material and instructional strategies, but not achieved the success demanded by NCLB. For example, teachers have taught “lowest common denominator” in the best way they know how, and some students have not learned it. Similarly, they have reviewed “noun-verb agreement” repeatedly, using different approaches,

and yet some students still make mistakes. If they spend more time on these topics, then they will have to omit other parts of the curriculum; if they press to cover all the content they will have to leave these skills unlearned. The problem for most educators is not that they are unaware of their students' failings and not that they are unmotivated to work harder, but that they have run out of options. Keep this picture in mind, while reading about quality improvement efforts in health care and manufacturing.

3. Quality Improvement in Health Care

As long as there have been physicians, there have been concerns about the quality of medical care. The earliest calls for quality improvement date back many centuries, and large-scale, formal efforts to measure and improve health care quality have been underway for the past 50 years (Lohr and Brook, 1984). Thus, health care provides a rich set of experiences for thinking about measuring and improving quality of service. The following paragraphs provide an abbreviated description of quality improvement efforts in health care over the past half-century.

There has been considerable debate among health researchers and physicians about the best way to define “quality,” but the framework that has been the most widely used is one that distinguishes between structure, process and outcomes (Donabedian, 1980). Structure refers to the “relatively stable characteristics of the providers of care, of the tools and resources they have at their disposal, and of the physical and organizational settings in which they work” (Donabedian, 1980, p. 81). The structural aspects of health care include the “bricks and mortar” of hospitals and clinics, the equipment that is available therein, and the underlying training and skills of the practitioners. Processes are the set of activities that occur within and between health practitioners (doctors, nurses, therapists, etc.) and patients. The quality of these processes can be judged against scientific, professional and social norms. Health outcomes are the “changes in a patient’s current and future health status that can be attributed to antecedent health care” (Donabedian, 1980, p. 83). Health status is usually defined broadly to include physiological, psychological and social functioning.

Ultimately, care is designed to improve health outcomes, and using any measure of quality other than outcomes is only justified if it is related to health. “If quality-of-care criteria based on structural or process data are to be credible, it must be demonstrated that variations in the attribute they measure lead to differences in outcome” (Brook, McGlynn and Cleary, 1996, p.966). There is evidence for both relationships, which is why all three types of measures have been used as indicators of health quality.

Measuring Quality Based on Structure

Structure was one of the first aspects of health quality to come under scrutiny. It is difficult to imagine effective medical care in the absence of an office, an x-ray machine, a stethoscope, antibiotics, or a certified physician. Early quality-improvement efforts focused on measuring the adequacy of the structural elements of health care available to clients in their communities. The structural elements of health care provide a minimum necessary basis for good care. When absent, for example, in a country with an emerging economy, then monitoring structural variables is probably the most important means for “protecting and promoting the quality of care” (Donabedian, 1990, p. 82).

In health care, the structural aspect of quality is embodied in accreditation agencies. The two major agencies are the Joint Commission on the Accreditation of Healthcare Organizations (JCAHO) and the National Committee for Quality Assurance (NCQA). Both began their work focusing on the structural aspects of quality, although both have broadened their reviews to

include more diverse measures. In recent years, JCAHO and NCQA data have become more public, and they are incorporated into report cards and on-line databases that make evidence about basic aspects of health providers more widely available.

The analogy to education is clear; accreditation has been one of the fundamental features of quality assurance in education for decades. Policymakers supported the creation of national, regional and state accreditation agencies covering all aspects of schooling from preschool to university. These agencies are designed to insure that each school has adequate buildings, libraries, textbooks, teachers and administrators. Without these basic foundational components, schools cannot provide quality education. In education, as in health care, structure is a necessary but not sufficient condition for quality.

Measuring Quality Based on Outcomes.

While structure serves as a reasonable starting point in efforts to insure quality health care, it is easy to see that good structure does not guarantee good care. As accreditation efforts grew more sophisticated, researchers and physicians concerned about quality improvement shifted their attention toward outcomes. Many argued that the time had come to stop focusing on antecedent conditions and start looking at direct measures of health. The key questions should be: Are patients better off after treatment than before? Which providers produce the best outcomes? As a result, quality improvement efforts began to focus on measuring and reporting outcomes.

The earliest public releases of data on health outcomes occurred in the 1980s, when the Health Care Financing Administration (HCFA) published hospital-specific mortality rates for Medicare patients. Some state health agencies began to follow suit. New York and Pennsylvania were among the leaders in releasing data on health outcomes. Since 1989, both states released data on the results of coronary arterial bypass graft (CABG) surgery performed at different hospitals (McGlynn and Brook, 2001). The New York Department of Health created the Cardiac Surgery Reporting System, which included severity-adjusted mortality rates by hospital, highlighting the difference between the institution and the state average, and identifying those hospitals and surgeons with rates higher or lower than the statewide rates (Chow and Stecher, 2003). Since the data were made public, risk-adjusted mortality rates for these procedures have fallen 41 percent (McGlynn and Brook, 2001).

Risk Adjustment Methods

One difficulty with judging quality in terms of outcomes is that all patients are not equal. "...outcomes are only partially produced by health services and are frequently influenced more by other factors, (e.g., natural history of the disease, patient physiologic reserve or patient age)" (Brook, McGlynn and Shekelle, 2000, p. 284). The recognition that outcomes can be a poor indicator of quality of care lead to the development of "risk adjustment" techniques to hold providers accountable only for their own caregiving actions and not for patient characteristics beyond their control. The goal of risk adjustment is to control for the effects of different initial patient characteristics when making provider-to-provider comparisons, thus "leveling the playing field" among providers (Iezzoni, 1994). In theory, risk adjustment can make it just as easy to demonstrate positive outcomes with a difficult-to-treat patient as with an easy-to-treat one. This

adjustment is important lest the measurements induce providers to reject sicker patients in favor of healthier ones with better outcomes.

A brief explanation of risk adjustment methods is sufficient for the purposes of this paper. Risk adjustment uses regression techniques to “control” for pre-existing differences in patient characteristics that might influence the outcomes of care. After deciding on an outcome to be studied, the analyst identifies the factors that independently affect this outcome and are not under the control of the provider (e.g., age, severity of coexisting illnesses). This is done on the basis of expert judgment and review of the scientific literature. Then a statistical model is developed with the outcome of interest as the dependent variable and the potential risk factors as the independent variables. Weights are derived from this model that can be assigned to each of the factors to predict the outcome for each patient regardless of the care received. The weights are applied to each patient seen by a given provider to predict the provider’s expected outcome rate. This rate is compared to the actual outcome rate as the basis for judging the quality of care provided. There are limitations to risk adjustment methods, and researchers are still working to develop more effective techniques. Yet, despite these limitations, risk adjustment methods are widely used in judging health care outcomes.

Many people believed that outcome data would influence the quality of care primarily by influencing consumer choice, i.e., patients would review provider data and select providers with better records. This would pressure providers to improve their patient outcomes. Surprisingly, the evidence suggests that “the information has only limited impact on consumer decision making” (Marshall, Shekelle et al., 2000, p. 1867). Rather, it appears that “provider organizations are more responsive to performance data than consumers, purchasers, or individual physicians” (Marshall, Shekelle et al., 2000, p. 1872). At present, report card-like information is largely ignored by patients and by physicians, while provider organizations, like hospitals or health plans, use the information to spur internal improvement efforts, and also use it as a marketing tool. Many believe this situation will continue. “We believe that the use of public performance data by consumers and purchasers or for regulation purposes will remain relatively less important for the foreseeable future than use of the data as a catalyst to stimulate and promote internal quality improvement mechanisms at the level of the organizational provider” (Marshall, Shekelle et al., 2000, p. 1874).

In the past decade, the provision of information on health provider performance has grown rapidly in the U.S., and the number of available “report cards,” “provider profiles,” or “consumer reports” has increased quickly. The production and dissemination of report cards is now a multimillion-dollar industry (Marshall, Shekelle et al., 2000). Researchers and policymakers hope that such data will serve many purposes, including promoting an efficient market economy for health care, encouraging consumers to choose high-quality providers, providing marketing evidence for providers, controlling costs, ensuring the accountability of providers, and promoting quality improvement. These gains have not been realized, yet.

Education is engaged in a flurry of outcome-based reporting that is similar to the situation that occurred in health care. The effort has encountered similar problems, and with similar effects (although it is too soon to make judgments about its long-term impact). One of the justifications for NCLB was to shift the focus of accountability from inputs (such as credentialing and

curriculum) to outcomes. Policymakers made sure that achievement was the paramount criterion of success under NCLB. However, just as all patients are not the same, all students are not the same. All students do not face similar challenges to become proficient. As a result, educational researchers are developing statistical techniques to level the playing field when making judgments about school quality. For example, value-added modeling is a currently popular approach to adjusting indicators to reflect the differences in students' initial status. Educators are also discovering that public reporting of results is not generating the market-based responses that were anticipated. Although it is too early to make any final judgments, few parents have opted to transfer their students to higher-performing schools. Yet, while parents are lukewarm in their response to outcome data, administrators are aggressively reviewing the test results and working hard to figure out how to improve results.

Measuring Quality Based on Process

In addition to the practice problems mentioned above, health researchers identified more fundamental problems with focusing quality improvement efforts on outcomes. Many patients get better even when they do not receive all the care they need, and some patients do worse even after receiving the best care. Furthermore, many deficiencies of care do not manifest themselves as medical problems for years (Brook, McGlynn and Cleary, 1996). For these reasons and others, processes are to be preferred over outcomes as a measure of quality, and current efforts in quality improvement in health care focus on process. This occurs even though process measures are less precise and more expensive to develop than measures of outcomes (Donabedian, 1980).

One of the reasons that process measures are more expensive to develop than outcome measures is because medical care is so complex and each encounter involves so many clinical choices. To many of us who have been patients, it might seem all but impossible to establish a set of procedural standards against which to judge each conceivable doctor-patient encounter. Fortunately, there is a huge body of research that establishes a sound basis for creating process measures in medicine. The basis is called "clinical practice guidelines."

Clinical Practice Guidelines¹

Guidelines are "systematically developed statements to assist practitioners' and patients' decisions about appropriate care for specified clinical outcomes" (IOM, 1990). They provide outlines of best practices for treating specific medical conditions. Independent organizations create the guidelines to support clinicians. While they are not mandatory, clinicians use them because they are based on scientific evidence and the judgment of experts.

Health guidelines come from a number of different sources. In 1989, the U.S. federal government created the Agency for Health Care Policy and Research, which is currently called the Agency for Health Care Research and Quality (AHRQ), to support the development of clinical practice guidelines and promote health outcomes research. The American Medical Association (AMA), other physician organizations, and medical specialty societies (e.g., the American Diabetes Association) also produce guidelines. Guidelines vary in format, but they are

¹ Parts of this discussion are adapted from Pearson and Stecher, 2003.

characterized by specific statements describing a recommended course of action for patients of a specific type under specific circumstances. The recommendations are often accompanied by an indication of the strength of the supporting evidence. Figure 1 shows an excerpt from one set of clinical practice guidelines for the treatment of diabetes.

Figure 3.1.
Prevention and Management of Diabetes Complications²

Blood Pressure Control

Screening and Diagnosis

* Blood pressure should be measured at every routine **diabetes** visit. Patients found to have systolic blood pressure >130 or diastolic blood pressure >80 mmHg should have blood pressure confirmed on a separate day. (C)

Goals

* Patients with **diabetes** should be treated to a systolic blood pressure <130 mmHg. (B)

* Patients with **diabetes** should be treated to a diastolic blood pressure <80 mmHg. (B)

Treatment

* Patients with hypertension (systolic blood pressure >140 or diastolic blood pressure >90 mmHg) should receive drug therapy in addition to lifestyle and behavioral therapy. (A)

* Multiple drug therapy (two or more agents at proper doses) is generally required to achieve blood pressure targets. (B)

* Patients with a systolic blood pressure of 130 to 139 mmHg or a diastolic blood pressure of 80 to 89 mmHg should be given lifestyle and behavioral therapy alone for a maximum of 3 months and then, if targets are not achieved, in addition, be treated with pharmacological agents that block the renin-angiotensin system. (E)

* Initial drug therapy for those with a blood pressure >140/90 mmHg should be with a drug class demonstrated to reduce cardiovascular disease (CVD) events in patients with **diabetes** (angiotensin converting enzymes [ACE] inhibitors, angiotensin receptor blockers [ARBs], beta-blockers, diuretics, and calcium channel blockers). (A)

* All patients with **diabetes** and hypertension should be treated with a regimen that includes either an ACE inhibitor or ARB. If one class is not tolerated, the other should be substituted. If needed to achieve blood pressure targets, a thiazide diuretic should be added. (E)

* If ACE inhibitors, ARBs, or diuretics are used, monitor renal function and serum potassium levels. (E)

Note: (A) – (E) reflect the quality of the evidence for the recommendation.

² Excerpt from *Standards of medical care in diabetes*. National Guideline Clearinghouse. http://www.guidelines.gov/summary/summary.aspx?doc_id=4679&nbr=3413&string=diabetes. Downloaded January 31, 2005.

Developing clinical practice guidelines involves four basic steps: (1) selection of an appropriate topic; (2) selection of the panel or group of people who will make the guideline decisions; (3) collection and presentation of the information on which the guideline decisions will be made; and (4) the decision-making process itself, which involves a variety of consensus building strategies like the Delphi technique.

The heart of guideline development is deciding which practices to recommend for given situations. Guideline developers must first identify the practices that produce the best outcomes, which are usually defined in terms of the greatest benefit to patients. The two most important questions are: “What are the best outcomes to strive toward?” and “Which processes produce these outcomes?” Appropriate practices are identified using both scientific evidence and expert opinion, with increasing emphasis on scientific evidence. Assessment of the scientific evidence involves extensive literature review, evaluation of the quality of the studies, and synthesis of the findings. Recommendations are usually graded or ranked by the strength of the findings. Although scientific research on the relationship between the processes and outcomes of health care has progressed rapidly in the past two decades, it still does not address the majority of clinical circumstance. Thus, expert opinion is an important means for identifying the practices which are appropriate for guideline recommendation.

There are approximately 2,000 guidelines available from the *Clinical Practice Guideline Directory*, maintained by the American Medical Association (AMA, 2000). And the National Guideline Clearinghouse (a joint effort of the AHRQ, the AMA and the American Association of Health Plans) maintains an online database and extensive information about clinical practice guidelines.

Guidelines can lead to improvements in care, but such changes are not automatic. In fact, research has shown limited impact of guideline dissemination on provider behavior in the absence of additional supports for change. However, there is strong evidence that guidelines, when effectively implemented, both change clinical practice and improve patient outcomes (Grimshaw, Freemantle et al., 1995). Guidelines can also help to inform patients of their options for treatment and the kinds of care to expect. A number of clinical practice guidelines—the AHCPR-sponsored guidelines, for example—include versions for the consumer as well as for the practitioner.

Quality Indicators

While clinical practice guidelines provide a detailed guide for practitioners, the level of detail is too great to use as part of a broad-based indicator system. One must narrow the focus in some manner to identify critical elements that can form the basis for summarizing the quality of care given to individual patients. The RAND Quality Assessment (QA) Tools is one approach to developing a manageable set of quality indicators of the “appropriateness of care” given to patients. The approach uses guidelines and expert judgment to focus on a subset of key procedures.

RAND staff selected 30 clinical areas representing the leading causes of death and disability as well as the major reasons that people seek care. They developed specific

standards or quality indicators within each clinical area, based on a review of the national guidelines and the medical literature. Panels of experts evaluated the proposed quality indicators.³ The researchers developed computer-assisted medical record abstraction software to facilitate data collection by nurses.

The researchers constructed a quality “score” for each patient. The score was the number of times that the patient received the care recommended across all of the conditions that the patient had in a two-year period, divided by the number of times that the patient was determined to need specific health care interventions. The team then aggregated the individual scores to obtain quality scores for various dimensions of performance (RAND Health, 2004, pp 1-2).

Figure 2 contains the Quality Indicator developed for Diabetes Mellitus.

Further research has confirmed the reliability of the panel rating process and the content and construct validity of the ratings of appropriateness (McGlynn and Brook, 2001). Many other researchers have used similar methods to evaluate the process quality of health care. The studies reveal both overuse and underuse of care. After reviewing a number of studies of the appropriateness of care, McGlynn and Brook (2001) report that overuse is high; they conclude, “about one-third of the procedures performed in this country are of questionable health benefit relative to their risks” (p. 165). At the same time, some patients do not receive the care they should. The National Report Card on Quality of Health Care in America found that “overall, adults received about half of the recommended care” (RAND Health, 2004, p. 1).

Quality Indicators are a relatively new development in health care quality improvement efforts, but they are already demonstrating the potential for impact. Indicators seem to be readily understood by the public and to command their attention. The recent RAND study was headline news in papers across the country. The *New York Times* reported, “Study finds widespread problem of inadequate health care” (May 5, 2004). The *Los Angeles Times* echoed, “Patients not getting the care they need...” (May 10, 2004).

Researchers are finding that they get better traction, in terms of quality improvement, from measures of practice than from measures of structure or outcomes. One important advantage of judging quality in terms of process is that the indicators identify specific targets for improvement.

Yet, quality indicator development is costly. “It takes skill, time and money to evaluate the scientific literature, update criteria as science changes, develop and administer valid data-collection instruments, and analyze the results with appropriate methods” (Brook, McGlynn and Cleary, 1996, p. 969). However, with support from government, foundations and from health plans, this approach to quality improvement seems to be gaining momentum.

³ All the quality indicators, including the literature reviewed and the panel recommendations, are available on the RAND Health website at www.rand.org/health/tools/qualist.html.

Figure 3.2.
Recommended Quality Indicators for Diabetes Mellitus

These indicators apply to men and women age 18 and older. Only the indicators in bold type were rated by this panel; the remaining indicators were endorsed by a prior panel.

Indicator	Quality of Evidence	Literature	Benefits	Comments
Diagnosis				
1. Patients with fasting blood sugar >126 or postprandial blood sugar >200 should have a diagnosis of diabetes noted in progress notes or problem list.	III	ADA, 1989 ECDCDM, 1997	Prevent diabetic complications. ¹	This definition of diabetes is accepted worldwide. Blood sugar tests are often ordered as part of panels.
2. Patients with the diagnosis of Type 1 diabetes should have all of the following: a. Glycosylated hemoglobin or fructosamine every 6 months. b. Eye and visual exam (annual). c. Total serum cholesterol and HDL cholesterol tests (annual). d. Measurement of urine protein (annual). e. Examination of feet at least twice a year. f. Measurement of blood pressure at every visit.	I, III	ADA, 1989; Larsen et al., 1990; ACP, ADA, and AAO, 1992	Prevent diabetic complications. ¹ Prevent retinopathy, hyperlipidemia, atherosclerotic complications, and renal disease.	Randomized controlled trial of 240 patients indicated a significant decrease in hemoglobin A1c among those whose hemoglobin A1c was monitored. Time interval is that used in most clinical trials. Eye and visual exams are shown to detect retinopathy at an earlier treatable stage. Other recommendations are based on expert opinion, though studies have shown conditions they screen for to be more common in diabetics and all are susceptible to treatment with improved outcomes resulting from earlier detection.
3. Patients with the diagnosis of Type 2 diabetes should have all of the following: a. Glycosylated hemoglobin or fructosamine every 6 months; b. Eye and visual exam (annual); c. Total serum cholesterol and HDL cholesterol tests (annual); d. Measurement of urine protein (annual); e. Examination of feet at least twice a year; f. Measurement of blood pressure at every visit.	I, III	ADA, 1989; Larsen et al., 1990; ACP, ADA, and AAO, 1992	Prevent diabetic complications. ¹ Prevent retinopathy, hyperlipidemia, atherosclerotic complications, and renal disease. Reduce morbidity from foot infections.	Randomized controlled trial of 240 patients indicated a significant decrease in hemoglobin A1c among those whose hemoglobin A1c was monitored. Time interval is that used in most clinical trials. Eye and visual exam are shown to detect retinopathy at an earlier treatable stage. Other recommendations are based on expert opinion, though studies have shown conditions they screen for to be more common in diabetics and all are susceptible to treatment with improved outcomes resulting from earlier detection.

Figure 3.2.
Recommended Quality Indicators for Diabetes Mellitus (cont.)

Indicator	Quality of Evidence	Literature	Benefits	Comments
4. Types 1 and 2 patients taking insulin should monitor their glucose at home unless documented to be unable or unwilling.	III	ADA, 1993	Prevent hypoglycemic episodes. Prevent diabetic complications.	A small RCT found that home glucose monitoring increases glycemic control in insulin-dependent diabetics. Another study found no difference in control by frequency of monitoring. Recommended by the ADA.
Treatment				
5. Newly diagnosed diabetics should receive dietary and exercise counseling.	II	Raz et al., 1994; Delahanty and Halford, 1993; ADA, 1989; Bergenstal et al., 1993	Reduce diabetic complications. ¹	Adherence to ADA diet decreases insulin and oral hypoglycemic requirements and serum lipids. Exercise improves glucose tolerance and may reduce or eliminate need for drug therapy. DCCT used dietitians and found that adherence to diet improved control, and the ADA and the ABFP recommend their use. No study has found that dietary counseling reduces diabetic complications.
6. Type 2 diabetics who have failed dietary therapy should receive oral hypoglycemic therapy.	III	ADA, 1989; Gerich, 1989; Bergenstal et al., 1993	Reduce diabetic complications. ¹	Observational trials have shown oral hypoglycemics to be effective in treating hyperglycemia and improving glycemic control. No studies have shown reduction of diabetic complications. Specialty societies and review articles widely recommend their use in mild to moderate disease before starting insulin.
7. Type 2 diabetics who have failed oral hypoglycemics should be offered insulin.	III	ADA, 1989; Bergenstal et al., 1993	Reduce diabetic complications.	Recommended by the ADA and ABFP.
8. Hypertensive diabetics with proteinuria should be offered an ACE inhibitor or a calcium channel blocker within 3 months of the notation of proteinuria.	I	Lederle, 1992; Anderson, 1990	Reduce diabetic complications.	May reduce progression to diabetic nephropathy.
9. All patients with diabetes should have a follow-up visit at least every 6 months.	III	Bergenstal et al., 1993; ADA, 1989	Reduce probability of severe diabetic complications.	Visits for diabetic patients in control should be every 3-6 months (per ABFP). Routine monitoring facilitates early detection and treatment of complications.

Source: Kerr, Asch, et al., (2002), pp. 141-143.

Figure 3.2.
Recommended Quality Indicators for Diabetes Mellitus (cont.)

Definitions and Examples

Diabetic complications include visual loss and dysfunction of the heart, peripheral vasculature, peripheral nerves, and kidneys. Synonyms for types 1 and 2 diabetes are listed below:

Type 1 diabetes	Type 2 diabetes
IDDM - Insulin-dependent diabetes Juvenile diabetes Juvenile-onset diabetes Ketosis-prone diabetes	AODM - Adult-onset diabetes MODM - Maturity-onset diabetes NIDDM - Non-insulin dependent diabetes mellitus Nonketosis-prone diabetes

Quality of Evidence Codes

I RCT
II-1 Nonrandomized controlled trials
II-2 Cohort or case analysis
II-3 Multiple time series
III Opinions or descriptive studies

4. Process Measurement Applied to Education

The conclusion I draw from this review of quality improvement in health care is that education should shift the focus of accountability from outcome quality to process quality. Outcome measures represented an advance over accreditation, but they suffer from the same fundamental flaws in education that they do in health care—students begin at different places and statistical “controls” are imperfect; some students learn despite poor instruction; some students fail to learn despite good instruction; and, most importantly, information about outcomes provides little guidance for improving processes.

Could we develop Quality Indicators for education? Not at present. Many of the preconditions that make process quality indicators possible in health care are absent in education. The essential preconditions include a body of scientific evidence about effective practice, clinical practice guidelines that establish standards for practice, and cumulative records of student status and services received. (A discussion of caveats at the end of the paper summarizes other differences.)

Despite these deficiencies, I believe the health care model offers a road map for a long-term development effort to create a process-based quality improvement system for education. This will not be accomplished overnight, and it is important to recognize that the health care quality improvement systems we now have reflect 50 years of research and development. Nevertheless, educators have developed structural quality measures and outcome quality measures, and there is no reason to believe that process quality measures cannot be created with appropriate commitment and support.

Briefly, here are the steps that would have to occur to reach a point equivalent to clinical practice guidelines in medicine.

1. Launch a formal, long-term effort to develop process quality measures for education. The launch will serve to attract talented researchers and financial support. In health care, the largest programs of quality assurance have been sponsored by the federal government, although there is growing participation of professional societies and provider organizations (Lohr and Brook, 1984). Federal support for similar efforts in education would be expected.
2. Develop a taxonomy of instructional “practice areas” to anchor the development effort. It seems clear to me that the fundamental events/interactions of practice should be embedded in subject matter, but I am not exactly sure of the correct size and scope. Potential starting points for this step are existing state standards, NAEP objectives, standards developed by the National Board for Professional Teaching Standards, and the work of the mathematics and reading study panels convened by RAND (Ball, 2003; Snow, 2002).
3. Select targets of opportunity, where scientific evidence and practical experience are most developed. Potential target “practice areas” include early literacy, elementary mathematics, or algebra.

4. Develop practice guidelines for the identified areas. This process includes: (a) assembling research evidence on effective practice; (b) convening panels of experts to review research; (c) categorizing research with respect to strength of evidence; (d) identifying gaps in the research; (e) soliciting best professional judgment to supplement research; (f), drafting practice guidelines; (g) distributing them for practitioner and researcher comment; and (h) revising and publishing the guidelines.
5. Initiate research to validate the guidelines and collect evidence to fill in gaps. This could include both formal scientific study and practitioner-based research.⁴
6. Disseminate the guidelines widely. Encourage their use in pre-service training, professional development, and performance review.

At the same time this effort is underway, we would need to initiate efforts to develop a student-level instructional record system. I believe this is a minor technological problem, but it has major work-flow consequences for teachers. If a student-level system proves unworkable in the short term, then a teacher-level system might be an acceptable alternative.

After the initial practice guidelines are produced, quality indicator development could begin. The process would be similar to the one outlined above for health care. The end result would be a procedure for rating schools on the basis of the instructional practices received by schools in key practice areas. They would permit summary judgments such as the following: Students at Borowitz School received 78 percent of the recommended reading instructional activities they should have received. They received 89 percent of the recommended mathematics learning activities. These statements would be followed by a list of specific practice areas in reading and mathematics, with tabulations for each, by grade level. It would also be possible to highlight services received by particular types of students, such as those who entered school with limited English language skills. Such reports would tell parents exactly what their children could expect if they attend Borowitz School. More importantly, they will provide educators at the school specific guidance about where their programs may be deficient and what steps they can take to improve the education they provide.

Caveats

Admittedly, this paper paints a rosy picture of educational quality improvement based on process indicators. Its purpose is to create a vision of an alternative to outcome-based accountability and, as a result it does not address every difficult issue. However, I acknowledge that there are weaknesses in the analogy between education and health care, and I have thought a lot about these differences. Some of the differences present serious obstacles to achieving this vision; others do not. Table 1 describes a number of the difference that might be relevant to the

⁴ The Toyota Production System included worker-initiated “hypothesis testing” for quality improvement that suggests an approach to school-based quality improvement in education. See Barney and Kirby (2004).

proposed approach to quality improvement. I have not dealt with all of them in this paper, but I have thought about them enough to believe that none presents an insurmountable obstacle.

Table 4.1.
Relevant Differences Between Health Care and Education

<i>Feature</i>	Health Care	Education
<i>Goal of Process</i>	Restore Health and Manage Chronic Disease	Promote Learning and Development
<i>Ultimate Measure of Success</i>	Patient Health Status	Student Achievement
<i>Singularity of Interaction</i>	One-on-One (sequential)	Many-to-One (simultaneous)
<i>Nature of Interaction</i>	Acute Event or Chronic Condition	Cumulative Development
<i>Length of Interaction</i>	Short, Episodic Contact	Continuous, Extended Contact
<i>Scope of Responsibility</i>	Broad	Broad
<i>Nature of Action</i>	Judgmental	Judgmental
<i>Degree of Openness</i>	Private	Public
<i>Support Infrastructure</i>	Huge, including Laboratories, Support Staff, etc.	Large, including Curriculum Materials, Aides, etc.
<i>Permanent Record</i>	Detailed Chart	General Cumulative File

References

- Adler, P. S. (1993, January-February). Time-and-motion regained. *Harvard Business Review*, 97-108.
- American Medical Association (2000). *Clinical practice guideline directory: 2000 edition*. Chicago, IL: American Medical Association.
- Ball, D. L. (Chair, RAND Mathematics Study Panel) (2003). *Mathematical proficiency for all students: Toward a strategic research and development program in mathematics education*. MR-1643-OERI. Santa Monica, CA: RAND.
- Barney, H. and Kirby, S. N. (2004). Toyota production system/lean manufacturing. In B. Stecher and S. N. Kirby (Eds.), *Organizational improvement and accountability. Lessons for education from other sectors*. MG-136-WFHF. Santa Monica, CA: RAND.
- Bauchner, H. and Simpson, L. (1998, October, Part II). Specific issues related to developing, disseminating, and implementing pediatric practice guidelines for physicians, patients, families, and other stakeholders. *Health Services Research*, 33(4), 1161-1177.
- Berwick, D. M., James, B., and Coye, M. J. (2003). Connections between quality measurement and improvement. *Medical Care*, 41(1) Supplement, I-30-I-38.
- Brook, R. H., McGlynn, E. A., and Cleary, P. D. (1996, September 26). Quality of health care. Part 2: Measuring quality of care. *New England Journal of Medicine*, 335(13), 966-970.
- Brook, R. H., McGlynn, E. A., and Shekelle, P. G. (2000). Defining and measuring quality of care: a perspective from U.S. researchers. *International Journal for Quality in Health Care*, 12(4), 281-295.
- California Department of Education (1998). *Steering by results: A high-stakes rewards and interventions program for California schools and students*. Sacramento, CA: Author.
- Chow, M. and Stecher, B. (2004). Health care report cards and the public release of data. In B. Stecher and S. N. Kirby (Eds.), *Organizational improvement and accountability. Lessons for education from other sectors*. MG-136-WFHF. Santa Monica, CA: RAND.
- Cook, C. R. and Graser, J. C. (2001). *Military airframe acquisition costs: The effects of lean manufacturing*. MR-1325-AF. Santa Monica, CA: RAND.
- Cook, C. R. (1996). Review of *Lean Work: Empowerment and exploitation in the global auto industry*, edited by Steve Babson. *Work and Occupations*, 23(3), 334-336.
- Donabedian, A. (1980). *Explorations in quality assessment and monitoring. Volume I: The definition of quality and approaches to its assessment*. Ann Arbor, MI: Health Administration Press.

Iezzoni, L. L. (Ed.) (1994). *Risk adjustment for measuring health care outcomes*. Ann Arbor, MI: Health Administration Press.

Institute of Medicine (1990). *Clinical practice guidelines: Directions for a new program*. Washington, DC: National Academy Press.

Kerr, E. A., Asch, S. M., Hamilton, E. G., and McGlynn, E. A. (Eds.) (2000). *Quality of care for general medical conditions. A review of the literature and quality indicators*. MR-1280-AHRQ. Santa Monica, CA: RAND.

Liker, J. K. (1998). Introduction: Bringing lean back to the USA. In J. K. Liker (Ed.), *Becoming lean: Inside stories of U.S. manufacturers*. Portland, OR: Productivity Press.

Lohr, K. N. and Brook, R. H. (1984, May/June). Quality assurance in medicine. *American Behavioral Scientist*, 27(5), 583-607.

Mangione-Smith, R. and McGlynn, E. A. (1998, October, Part II). Assessing the quality of healthcare provided to children. *Health Services Research*, 33(4), 1059-1090.

Marshall, M. N., Shekelle, P. G., Leatherman, S., and Brook, R. H. (2000, April 12). The public release of performance data: What do we expect to gain? A review of the evidence. *Journal of the American Medical Association*, 283(14): 1866-1874.

McGlynn, E. A. and Brook, R. H. (2001). Evaluating the quality of care. In R. M. Andersen, T. H. Rice, and G. F. Kominski (Eds.), *Changing the U.S. health care system: Key issues in health services policy and management*, 2nd edition. New York, NY: Jossey-Bass, Inc.

McGlynn, E. A. and Adams, J. L. (2001). Public release of information on quality. In R. M. Andersen, T. H. Rice, and G. F. Kominski (Eds.), *Changing the U.S. health care system: Key issues in health services policy and management*, 2nd edition. New York, NY: Jossey-Bass, Inc.

Pearson, M. and Stecher, B. (2004). Clinical practice guidelines in the health sector. In B. Stecher and S. N. Kirby (Eds.), *Organizational improvement and accountability. Lessons for education from other sectors*. MG-136-WFHF. Santa Monica, CA: RAND.

RAND Health (2004). *Research highlights. The first national report card on quality of health care in America*. Santa Monica, CA: Author.

Snow, C. E. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. MR-1465-OERI. Santa Monica, CA: RAND.

Spear, S. and Bowen, R. K. (1999, September-October). Decoding the DNA of the Toyota production system. Reprint 99509. *Harvard Business Review*, 97-106.